# GENETIC POLYMORPHISM IN CASEINS OF COW'S MILK. VI. AMINO ACID COMPOSITION OF $\alpha_{s1}$-CASEINS A, B, AND C

W. G. GORDON, J. J. BASCH, AND M. P. THOMPSON

Eastern Regional Research Laboratory,[1] USDA, Philadelphia, Pennsylvania

ABSTRACT

The amino acid composition of highly purified preparations of the genetically controlled variants of $\alpha_{s1}$-casein, A, B, and C has been determined. B and C are virtually identical in composition, the only significant difference being that C contains one more glycine residue per mole than B. The data suggest that in B the missing glycine might be replaced by glutamic acid. The differences between A and the other variants are numerous and more substantial and, at present, cannot be accounted for in terms of single amino acid substitutions.

The genetic control of variants of $\alpha_{s1}$-caseins and procedures for the isolation and purification of the proteins have been subjects of earlier papers in this series (10, 13). We wish now to describe the amino acid composition of highly purified $\alpha_{s1}$-caseins A, B, and C. Preliminary communications (2, 14), in which differences in amino acid composition among the variants were discussed, were based on results obtained on unchromatographed, less pure preparations, and some of the differences previously reported may be attributed to the presence of small amounts of contaminants in the proteins analyzed.

## EXPERIMENTAL PROCEDURE

$\alpha_{s1}$-Caseins A, B, and C were isolated from milks of individual, typed cows, then purified by twice chromatographing on DEAE-cellulose in the presence of 3.3 M urea, and prepared for analysis by techniques already described (13); some chemical and physical characteristics of these preparations have also been reported (13).

Weighed samples (about 2 mg) of the proteins were hydrolyzed in 1-ml portions of glass-distilled 6 N HCl in sealed, evacuated tubes maintained at 110 ± 1 C in a forced-draft oven for periods of 24, 72, and 96 hr. Usually three (in a few cases, four) samples of each protein were hydrolyzed at each period, after which the tubes were opened, excess HCl removed by means of a rotary evaporator, and the residual amino acids dissolved in a known volume of pH 2.2 buffer containing 5 ml per liter of thiodiglycol. Aliquots were then analyzed on the 15- and 150-cm columns of an automatic recording analyzer (11).

Tryptophan was determined in replicate samples of the solid proteins by procedure "N" of Spies and Chambers (12); the time of Reaction I was 6.5 hr, and of Reaction II, 0.5 hr.

## RESULTS AND DISCUSSION

The analytical results are summarized in Table 1. Values shown are averages of analyses of nine or ten hydrolyzates, with the following exceptions: Results for serine and threonine are extrapolated values by the method of least squares, since progressive decomposition of these amino acids occurred; similarly, from the ammonia values, which increase with time, approximate values for amide ammonia were obtained by extrapolation to zero time. Determinations of valine, isoleucine, and histidine in 24-hr hydrolyzates were omitted in the calculation of final averages because of incomplete liberation of these amino acids. It may be pointed out that incomplete liberation of histidine in 24 hr is unusual, but it was evident in the analyses of all three variants. The tryptophan values, as noted in the table, are averages of four replicate analyses.

The tabulated results show, in general, many similarities in amino acid composition, as may be expected among genetic variants of $\alpha_{s1}$-caseins. The presence of cystine or cysteine could not be detected in any of the chromatograms of $\alpha_{s1}$-caseins, an observation in agreement with the previous findings of Waugh et al. (15) with $\alpha_{s1,2}$-caseins. Comparison of the results for the B and C variants reveals only a single difference in composition; namely, in glycine content, which is statistically significant. However, when the A and B variants are compared, many significant differences are apparent, the most pro-

[1] Eastern Utilization Research and Development Division, Agricultural Research Service, United States Department of Agriculture.

TABLE 1

Amino acid composition of $\alpha_{s1}$-caseins
(Grams amino acid per 100 g dry protein)

| | A | B | C |
|---|---|---|---|
| Aspartic Acid | 7.98 ± 0.12 | 8.42 ± 0.12 | 8.47 ± 0.14 |
| Threonine | 2.84 ± 0.15 | 2.51 ± 0.11 | 2.54 ± 0.08 |
| Serine | 6.65 ± 0.31 | 6.36 ± 0.36 | 6.45 ± 0.22 |
| Glutamic Acid | 24.46 ± 0.42 | 23.88 ± 0.38 | 23.39 ± 0.32 |
| Proline | 8.47 ± 0.15 | 8.19 ± 0.22 | 8.20 ± 0.15 |
| Glycine | 2.86 ± 0.04 | 2.82 ± 0.04 | 3.11 ± 0.06 |
| Alanine | 3.16 ± 0.04 | 3.35 ± 0.06 | 3.38 ± 0.10 |
| Valine | 4.99 ± 0.13 | 5.49 ± 0.12 | 5.58 ± 0.23 |
| Methionine | 3.12 ± 0.08 | 2.98 ± 0.12 | 2.96 ± 0.06 |
| Isoleucine | 6.38 ± 0.10 | 5.99 ± 0.10 | 6.10 ± 0.19 |
| Leucine | 8.11 ± 0.14 | 9.31 ± 0.11 | 9.41 ± 0.21 |
| Tyrosine | 7.80 ± 0.12 | 7.36 ± 0.14 | 7.41 ± 0.18 |
| Phenylalanine | 4.41 ± 0.09 | 5.52 ± 0.09 | 5.60 ± 0.13 |
| Lysine | 9.47 ± 0.18 | 8.71 ± 0.14 | 8.71 ± 0.15 |
| Histidine | 3.44 ± 0.10 | 3.33 ± 0.09 | 3.32 ± 0.06 |
| Amide ammonia | 1.65 ± 0.31 | 1.85 ± 0.25 | 1.76 ± 0.14 |
| Arginine | 3.88 ± 0.10 | 4.37 ± 0.06 | 4.38 ± 0.09 |
| Tryptophan | 2.01 | 2.03 | 2.00 |

Figures are average or extrapolated values (see text) with $\pm t_{.05}s\bar{x}$, the 95% confidence limits, as measure of repeatability; this means that if the experiment were repeated, the new values would be within the values shown ± the 95% confidence limits, 95% of the time. The tryptophan figures are simple averages of four analyses.

nounced being those in content of leucine and phenylalanine. Smaller differences in aspartic acid, threonine, alanine, valine, isoleucine, tyrosine, lysine, and arginine content are also evident.

Perhaps a more meaningful comparison of the composition of the variants may be obtained from the same results when they are calculated in terms of residues of amino acid per mole of protein. To determine molecular weights from the analytical results, the simple calculation of minimum molecular weights from percentages of amino acids present in smallest amounts was made. Tryptophan, although present in lowest concentration, was not selected for these calculations because the method of analysis leaves something to be desired with certain proteins; for example, $\beta$-lactoglobulin and $\alpha$-lactalbumin (3, 4). Of the other amino acids only threonine, methionine, histidine, and arginine are found in amounts low enough, in terms of numbers of residues, for the calculations to be useful. Estimates of molecular weights derived in this way are shown in Table 2. They agree reasonably well with the value of 27,300, based on light-scattering studies, reported for $\alpha_{s1,2}$-caseins by Dreizen, Noble, and Waugh (1). Unpublished results in this laboratory obtained by Noelken and Swaisgood for $\alpha_{s1}$-B and C by physical methods indicate that the molecular weights of these variants are in the range of 28,000-30,000. Kalan, Thompson, and Greenberg (9) reported molecular weights of about 26,000 for A and 31,000 for B and C, calculated from the amount of C-terminal tryptophan released by carboxypeptidase A.

Whether the somewhat smaller molecular weight of 28,000 for $\alpha_{s1}$-casein A in Table 2 is significantly different from the value of 28,600 found for B and C is doubtful, but we have used these values to calculate the number of amino acid residues per mole of protein. Results are listed in Table 3. It is obvious that the absolute numbers of amino acid residues depend on the molecular weights chosen. With this reservation in mind, it may be seen that the figures for B and C may be rounded off fairly well to whole numbers. Clearly, B and C differ in content of glycine by one residue. Since we are dealing here with genetic variants of the same protein, it is most probable that the missing glycine residue in B is replaced by some other additional amino acid residue. Our data suggest that this might be glutamic acid. The substitution, glycine for glutamic acid, is not an uncommon one; it has been demonstrated in a number of protein polymorphism studies (8). Admittedly, however, the analytical data do not provide convincing evidence for the glutamic acid difference and proof may well depend on isolation and analysis of difference peptides, which have already been visualized in experiments on the fingerprinting of enzymatic digests of the $\alpha_{s1}$-caseins.

Comparison of variant A with either B or C with respect to the occurrence of possible single amino acid replacements is manifestly fruitless

TABLE 2

Estimation of molecular weights of $\alpha_{s1}$-caseins from amino acid analyses

| | A | | | | B | | | | C | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | % Amino acid | Min mol wt | Assumed no. residues | Calc'd mol wt | % Amino acid | Min mol wt | Assumed no. residues | Calc'd mol wt | % Amino acid | Min mol wt | Assumed no. residues | Calc'd mol wt |
| Threonine | 2.84 | 4,194 | 7 | 29,358 | 2.50 | 4,765 | 6 | 28,590 | 2.53 | 4,708 | 6 | 28,248 |
| Methionine | 3.12 | 4,782 | 6 | 28,692 | 2.98 | 5,007 | 6 | 30,042 | 2.96 | 5,041 | 6 | 30,246 |
| Histidine | 3.44 | 4,510 | 6 | 27,060 | 3.33 | 4,659 | 6 | 27,954 | 3.32 | 4,673 | 6 | 28,038 |
| Arginine | 3.88 | 4,490 | 6 | 26,940 | 4.37 | 3,986 | 7 | 27,902 | 4.38 | 3,977 | 7 | 27,839 |
| Average | | | | 28,013 | | | | 28,622 | | | | 28,593 |

TABLE 3

Amino acid composition of $\alpha_{s1}$-caseins (Residues amino acid per mole protein)

| | A mol wt = 28,000 | B mol wt = 28,600 | C mol wt = 28,600 |
|---|---|---|---|
| Aspartic Acid | 16.8 | 18.1 | 18.2 |
| Threonine | 6.7 | 6.0 | 6.1 |
| Serine | 17.8 | 17.3 | 17.6 |
| Glutamic Acid | 46.6 | 46.4 | 45.5 |
| Proline | 20.6 | 20.3 | 20.4 |
| Glycine | 10.7 | 10.7 | 11.8 |
| Alanine | 9.9 | 10.8 | 10.8 |
| Valine | 11.9 | 13.4 | 13.6 |
| Methionine | 5.9 | 5.7 | 5.7 |
| Isoleucine | 13.6 | 13.1 | 13.3 |
| Leucine | 17.3 | 20.3 | 20.5 |
| Tyrosine | 12.1 | 11.6 | 11.7 |
| Phenylalanine | 7.5 | 9.6 | 9.7 |
| Lysine | 18.1 | 17.0 | 17.0 |
| Histidine | 6.2 | 6.1 | 6.1 |
| Amide ammonia | 27.1 | 31.1 | 29.7 |
| Arginine | 6.2 | 7.2 | 7.2 |
| Tryptophan | 2.8 | 2.7 | 2.8 |

at this time. Variant A appears to have not only three fewer leucines and two fewer phenylalanines, but also one less alanine, one less valine, and one less arginine than B or C, and one less glycine than C. The deficiencies are only partially offset by the presence of one more threonine, one more lysine, and possibly one more each of isoleucine and tyrosine. These differences are also detectable in Table 1, where equal weights of protein are being compared. Yet, $\alpha_{s1}$-caseins A and B are very much alike in content of seven amino acids and in lack of cystine, as well as in a number of other chemical and physical properties previously described (13). Furthermore, many identical spots in the fingerprinting experiments currently in progress support the inference that considerable portions of the molecules of $\alpha_{s1}$-caseins must be quite similar.

As we have pointed out previously (10), $\alpha_{s1}$-casein A has been observed only in the milks of a single blood line of Holstein cows. It has never been observed in genetic surveys in England, France, or the Netherlands. The age of this unique $\alpha_{s1}$-casein A variant is open to much speculation. Is it a mutant which has been dying out and is being revived by artificial breeding or is it a mutant of more recent origin? The manifold differences in amino acid composition strongly suggest that it is not a mutant of B and cause us to ask, "Will other casein variants of the $\alpha_{s1}$-series yet be discovered?"

The essential completeness of the amino acid analyses in Table 1 may be shown by calculation of per cent recoveries of nitrogen and of weight. Based on the figures of 15.10, 15.34, and 15.40% nitrogen in $\alpha_{s1}$-caseins A, B, and C, respectively (13), recoveries in terms of amino acid plus amide nitrogen were 101.6, 101.6, and 102.2%. In terms of weight, the recoveries were 97.8, 98.3, and 98.6%, including phosphoric acid residues.

Finally, it is of interest to compare the present results for $\alpha_{s1}$-caseins with those previously published for some related components of casein. For $\alpha_{s1,2}$-caseins, Waugh et al. (15) reported the presence of 2.3 tryptophans and 11 tyrosines per 27,300 g and, more recently, Ho and Waugh (6) completed the analysis of this preparation; for the majority of amino acids our values agree fairly well, but we are at a loss to account for the very serious discrepancies in serine, proline, aspartic acid, and amide ammonia content. Other comparisons may be made from the data in Table 4, where the

TABLE 4

Comparison of amino acid composition of some caseins
(Grams amino acid per 100 g protein)

| | $\alpha$-Casein (5) | $\alpha_{s1}$-Casein BC [a] | $\kappa$-Casein (7) | $\alpha_3$-Casein (5) |
|---|---|---|---|---|
| Aspartic Acid | 8.05 | 8.4 | 7.30 | 7.59 |
| Threonine | 4.2 | 2.5 | 6.64 | 4.2 |
| Serine | 6.4 | 6.4 | 6.09 | 5.7 |
| Glutamic Acid | 21.0 | 23.6 | 17.35 | 18.0 |
| Proline | 8.24 | 8.2 | 8.78 | 10.3 |
| Glycine | 2.01 | 3.0 | 1.31 | 1.27 |
| Alanine | 3.48 | 3.4 | 5.41 | 5.54 |
| Half-cystine | 0.44 | 0.0 | 1.40 | 1.34 |
| Valine | 5.70 | 5.6 | 5.10 | 5.15 |
| Methionine | 2.67 | 3.0 | 1.0 | 1.14 |
| Isoleucine | 5.75 | 6.0 | 6.14 | 6.45 |
| Leucine | 7.98 | 9.4 | 6.08 | 6.46 |
| Tyrosine | 7.30 | 7.4 | 7.40 | 9.80 |
| Phenylalanine | 4.66 | 5.6 | 4.07 | 4.00 |
| Lysine | 9.31 | 8.7 | 5.76 | 6.79 |
| Histidine | 2.92 | 3.3 | 1.67 | 1.51 |
| Arginine | 3.92 | 4.4 | 4.0 | 4.53 |
| Tryptophan | 2.00 | 2.0 | 1.05 | 1.82 |

[a] Figures in this column are average values from the analyses in Table 1 of a 1:1 mixture of the B and C variants.

amino acid composition of $\alpha$-casein, $\alpha_{s1}$-casein BC, $\kappa$-casein, and $\alpha_3$-casein are listed. The fact that $\alpha$-casein is made up of a mixture of $\alpha_{s1}$- and $\kappa$- (or $\alpha_3$-) caseins is readily apparent. With few exceptions, when the figure for $\alpha_{s1}$-casein is higher than that for $\alpha$-casein, the figure for $\kappa$-casein is lower, and vice versa.

## ACKNOWLEDGMENT

We thank Sandra R. Hoffer for her excellent operation of the amino acid analyzer and Joe N. Boyd, Biometrical Services, for statistical treatment of the data.

## REFERENCES

(1) Dreizen, P., Noble, R. W., and Waugh, D. F. 1962. Light Scattering Studies of $\alpha_{s1,2}$-Caseins. J. Am. Chem. Soc., 84: 4938.

(2) Gordon, W. G., and Basch, J. J. 1963. Differences in Amino Acid Composition of Genetic Variants of Bovine $\alpha_s$-Caseins. Federation Proc., 22: 657.

(3) Gordon, W. G., Basch, J. J., and Kalan, E. B. 1961. Amino Acid Composition of $\beta$-Lactoglobulins A, B, and AB. J. Biol. Chem., 236: 2908.

(4) Gordon, W. G., and Semmett, W. F. 1953. Isolation of Crystalline $\alpha$-Lactalbumin from Milk. J. Am. Chem. Soc., 75: 328.

(5) Hipp, N. J., Basch, J. J., and Gordon, W. G. 1961. Amino Acid Composition of $\alpha_1$-, $\alpha_2$-, and $\alpha_3$-Caseins. Arch. Biochem. Biophys., 94: 35.

(6) Ho, C., and Waugh, D. F. 1965. Interactions of Bovine $\alpha_s$-Caseins with Small Ions. J. Am. Chem. Soc., 87: 110.

(7) Jollès, P., Alais, C., and Jollès, J. 1962. Amino Acid Composition of $\kappa$-Casein and Terminal Amino Acids of $\kappa$- and Para-$\kappa$-Casein. Arch. Biochem. Biophys., 98: 56.

(8) Jukes, T. H. 1962. Relations Between Mutations and Base Sequences in the Amino Acid Code. Proc. Natl. Acad. Sci., 48: 1809.

(9) Kalan, E. B., Thompson, M. P., and Greenberg, Rae. 1964. End-Group Analysis of $\alpha_{s1}$-Caseins A, B, and C. Arch. Biochem. Biophys., 107: 521.

(10) Kiddy, C. A., Johnston, J. O., and Thompson, M. P. 1964. Genetic Polymorphism in Caseins of Cow's Milk. I. Genetic Control of $\alpha_s$-Casein Variation. J. Dairy Sci., 47: 147.

(11) Moore, S., and Stein, W. H. 1963. Chromatographic Determination of Amino Acids by the Use of Automatic Recording Equipment. Vol. VI, p. 819. S. P. Colowick and N. O. Kaplan, eds. Academic Press, New York.

(12) Spies, J. R., and Chambers, D. C. 1949. Chemical Determination of Tryptophan in Proteins. Anal. Chem., 21: 1249.

(13) Thompson, M. P., and Kiddy, C. A. 1964. Genetic Polymorphism in Caseins of Cow's Milk. III. Isolation and Properties of $\alpha_{s1}$-Caseins A, B, and C. J. Dairy Sci., 47: 626.

(14) Thompson, M. P., Pepper, L., Gordon, W. G., and Basch, J. J. 1963. The Chemical Composition of Bovine $\alpha_s$-A, B, and C Caseins. J. Dairy Sci., 46: 607.

(15) Waugh, D. F., Ludwig, M. L., Gillespie, J. M., Melton, B., Foley, M., and Kleiner, E. S. 1962. The $\alpha_s$-Caseins of Bovine Milk. J. Am. Chem. Soc., 84: 4929.